#17

1632 1600

RAW SEQUENCE LISTING DATE: 11/21/2002
PATENT APPLICATION: US/09/653,812B TIME: 11:04:50

Input Set : A:\09596-23U3 (053893-5006-03).txt
Output Set: N:\CRF4\11212002\I653812B.raw

3 <110> APPLICANT: KAZAZIAN, HAIG H.
4 BOEKE, JEF D.
5 MORAN, JOHN V.
6 BETH, DOMBROSKI A.
8 <120> TITLE OF INVENTION: COMPOSITIONS AND METHODS OF USE OF MAMMALIAN RETROTRANSPOSONS
10 <130> FILE REFERENCE: 9596-23U3 (053893-5006-02)
12 <140> CURRENT APPLICATION NUMBER: 09/653,812B
13 <141> CURRENT FILING DATE: 2000-09-01
15 <150> PRIOR APPLICATION NUMBER: US 08/847,844
16 <151> PRIOR FILING DATE: 1997-04-28
18 <150> PRIOR APPLICATION NUMBER: US 08/749,805
19 <151> PRIOR FILING DATE: 1996-11-15
21 <150> PRIOR APPLICATION NUMBER: US 60/006,831
22 <151> PRIOR FILING DATE: 1995-11-16
24 <160> NUMBER OF SEQ ID NOS: 152
26 <170> SOFTWARE: PatentIn version 3.1

28 <210> SEQ ID NO: 1
29 <211> LENGTH: 14
30 <212> TYPE: PRT
31 <213> ORGANISM: Neurospora crassa
33 <400> SEQUENCE: 1
35 Met Val Gln Leu Lys Ile Leu Tyr Trp Asn Val Gly Lys Ser
36 1 5 10
39 <210> SEQ ID NO: 2
40 <211> LENGTH: 10
41 <212> TYPE: PRT
42 <213> ORGANISM: Neurospora crassa
44 <400> SEQUENCE: 2
46 Tyr Asp Ile Val Ala Ile Gln Glu Pro Gly
47 1 5 10
50 <210> SEQ ID NO: 3
51 <211> LENGTH: 10
52 <212> TYPE: PRT
53 <213> ORGANISM: Neurospora crassa
55 <400> SEQUENCE: 3
57 Lys Gly Arg Ala Val Ile Tyr Val Asn Lys
58 1 5 10
61 <210> SEQ ID NO: 4
62 <211> LENGTH: 13
63 <212> TYPE: PRT
64 <213> ORGANISM: Neurospora crassa
66 <400> SEQUENCE: 4
68 Pro Thr Thr Val Tyr Ser Ile Tyr Ser Pro Ile Leu Thr

RAW SEQUENCE LISTING DATE: 11/21/2002
PATENT APPLICATION: **US/09/653,812B** TIME: 11:04:50

Input Set : **A:\09596-23U3 (053893-5006-03).txt**
Output Set: **N:\CRF4\11212002\I653812B.raw**

```
69 1               5                   10
72 <210> SEQ ID NO: 5
73 <211> LENGTH: 16
74 <212> TYPE: PRT
75 <213> ORGANISM: Neurospora crassa
77 <400> SEQUENCE: 5
79 Asn Leu Val Ala Val Gly Asp Leu Asn Leu His His Pro Asp Trp Asp
80 1               5                   10                  15
83 <210> SEQ ID NO: 6
84 <211> LENGTH: 16
85 <212> TYPE: PRT
86 <213> ORGANISM: Neurospora crassa
88 <400> SEQUENCE: 6
90 Gly Glu Pro Thr Arg Leu Gly Asn Ala Thr Arg Gly Glu Arg Asp Gly
91 1               5                   10                  15
94 <210> SEQ ID NO: 7
95 <211> LENGTH: 13
96 <212> TYPE: PRT
97 <213> ORGANISM: Neurospora crassa
99 <400> SEQUENCE: 7
101 Gly Ser Asp His Cys Pro Gln Glu Ile Trp Val Gln Val
102 1               5                   10
105 <210> SEQ ID NO: 8
106 <211> LENGTH: 19
107 <212> TYPE: PRT
108 <213> ORGANISM: Trypanosoma cruzi
110 <400> SEQUENCE: 8
112 Asp Ile Glu Gln Asn Pro Gly Pro Ile Ala Val Leu Gln Met Asn Val
113 1               5                   10                  15
116 Ser Cys Leu
120 <210> SEQ ID NO: 9
121 <211> LENGTH: 10
122 <212> TYPE: PRT
123 <213> ORGANISM: Trypanosoma cruzi
125 <400> SEQUENCE: 9
127 Ala Asp Ile Ile Ala Ile Gln Glu Thr Trp
128 1               5                   10
131 <210> SEQ ID NO: 10
132 <211> LENGTH: 10
133 <212> TYPE: PRT
134 <213> ORGANISM: Trypanosoma cruzi
136 <400> SEQUENCE: 10
138 Gly Gly Gly Val Ala Val Leu Val Arg Lys
139 1               5                   10
142 <210> SEQ ID NO: 11
143 <211> LENGTH: 13
144 <212> TYPE: PRT
145 <213> ORGANISM: Trypanosoma cruzi
147 <400> SEQUENCE: 11
```

RAW SEQUENCE LISTING DATE: 11/21/2002
PATENT APPLICATION: **US/09/653,812B** TIME: 11:04:50

Input Set : **A:\09596-23U3 (053893-5006-03).txt**
Output Set: **N:\CRF4\11212002\I653812B.raw**

```
149 Asp Leu Ile Val Ala Ser Ala Tyr Met Arg Pro Pro Pro
150 1               5                   10
153 <210> SEQ ID NO: 12
154 <211> LENGTH: 16
155 <212> TYPE: PRT
156 <213> ORGANISM: Trypanosoma cruzi
158 <400> SEQUENCE: 12
160 Pro Leu Leu Leu Cys Gly Asp Phe Asn Met His His Pro Gln Trp Glu
161 1               5                   10                  15
164 <210> SEQ ID NO: 13
165 <211> LENGTH: 21
166 <212> TYPE: PRT
167 <213> ORGANISM: Trypanosoma cruzi
169 <400> SEQUENCE: 13
171 Gly Glu Ile Thr Thr Ala Arg Gly Thr Arg Glu Arg Ser Cys Ile Asp
172 1               5                   10                  15
175 Leu Thr Trp Ser Lys
176             20
179 <210> SEQ ID NO: 14
180 <211> LENGTH: 13
181 <212> TYPE: PRT
182 <213> ORGANISM: Trypanosoma cruzi
184 <400> SEQUENCE: 14
186 Leu Ser Asp His Tyr Val Leu Thr Phe Thr Leu His Gln
187 1               5                   10
190 <210> SEQ ID NO: 15
191 <211> LENGTH: 17
192 <212> TYPE: PRT
193 <213> ORGANISM: Bombyx mori
195 <400> SEQUENCE: 15
197 Met Asp Ile Arg Pro Arg Leu Pro Ile Gly Gln Ile Asn Leu Gly Gly
198 1               5                   10                  15
201 Ala
205 <210> SEQ ID NO: 16
206 <211> LENGTH: 10
207 <212> TYPE: PRT
208 <213> ORGANISM: Bombyx mori
210 <400> SEQUENCE: 16
212 Leu Asp Ile Val Leu Val Gln Glu Gln Tyr
213 1               5                   10
216 <210> SEQ ID NO: 17
217 <211> LENGTH: 10
218 <212> TYPE: PRT
219 <213> ORGANISM: Bombyx mori
221 <400> SEQUENCE: 17
223 Lys Ala Gly Val Tyr Ile Arg Asn Arg Val
224 1               5                   10
227 <210> SEQ ID NO: 18
228 <211> LENGTH: 13
```

RAW SEQUENCE LISTING DATE: 11/21/2002
PATENT APPLICATION: US/09/653,812B TIME: 11:04:50

Input Set : A:\09596-23U3 (053893-5006-03).txt
Output Set: N:\CRF4\11212002\I653812B.raw

```
229 <212> TYPE: PRT
230 <213> ORGANISM: Bombyx mori
232 <400> SEQUENCE: 18
234 Asp Leu Tyr Met Val Ser Ala Tyr Phe Gln Tyr Ser Asp
235 1               5                   10
238 <210> SEQ ID NO: 19
239 <211> LENGTH: 16
240 <212> TYPE: PRT
241 <213> ORGANISM: Bombyx mori
243 <400> SEQUENCE: 19
245 Arg Val Val Ile Cys Ala Asp Thr Asn Ala His Ser Pro Leu Trp His
246 1               5                   10                  15
249 <210> SEQ ID NO: 20
250 <211> LENGTH: 21
251 <212> TYPE: PRT
252 <213> ORGANISM: Bombyx mori
254 <400> SEQUENCE: 20
256 Gly His Leu Pro Thr Phe Ser Thr Ala Asn Gly Glu Ser Tyr Val Asp
257 1               5                   10                  15
260 Val Thr Leu Ser Thr
261             20
264 <210> SEQ ID NO: 21
265 <211> LENGTH: 13
266 <212> TYPE: PRT
267 <213> ORGANISM: Bombyx mori
269 <400> SEQUENCE: 21
271 Ser Ser Asp His Arg Leu Ile Val Phe Gly Val Gly Gly
272 1               5                   10
275 <210> SEQ ID NO: 22
276 <211> LENGTH: 15
277 <212> TYPE: PRT
278 <213> ORGANISM: Drosophila melanogaster
280 <400> SEQUENCE: 22
282 Ile Met Ala Thr Leu Phe Ile Ala Thr Trp Asn Ala Asn Gly Val
283 1               5                   10                  15
286 <210> SEQ ID NO: 23
287 <211> LENGTH: 10
288 <212> TYPE: PRT
289 <213> ORGANISM: Drosophila melanogaster
291 <400> SEQUENCE: 23
293 Ile Asp Val Met Leu Leu Ser Glu Thr His
294 1               5                   10
297 <210> SEQ ID NO: 24
298 <211> LENGTH: 10
299 <212> TYPE: PRT
300 <213> ORGANISM: Drosophila melanogaster
302 <400> SEQUENCE: 24
304 His Gly Gly Thr Ala Ile Leu Ile Arg Asn
305 1               5                   10
```

RAW SEQUENCE LISTING DATE: 11/21/2002
PATENT APPLICATION: US/09/653,812B TIME: 11:04:50

Input Set : A:\09596-23U3 (053893-5006-03).txt
Output Set: N:\CRF4\11212002\I653812B.raw

```
308 <210> SEQ ID NO: 25
309 <211> LENGTH: 13
310 <212> TYPE: PRT
311 <213> ORGANISM: Drosophila melanogaster
313 <400> SEQUENCE: 25
315 Leu Leu Thr Leu Ala Ala Val Tyr Cys Pro Pro Arg Phe
316 1               5                   10
319 <210> SEQ ID NO: 26
320 <211> LENGTH: 16
321 <212> TYPE: PRT
322 <213> ORGANISM: Drosophila melanogaster
324 <400> SEQUENCE: 26
326 His Phe Ile Ala Ala Gly Asp Tyr Asn Ala Lys His Thr His Trp Gly
327 1               5                   10                  15
330 <210> SEQ ID NO: 27
331 <211> LENGTH: 24
332 <212> TYPE: PRT
333 <213> ORGANISM: Drosophila melanogaster
335 <400> SEQUENCE: 27
337 Pro Gly Ser Pro Thr Tyr Trp Pro Ser Asp Leu Asn Lys Leu Pro Asp
338 1               5                   10                  15
341 Leu Ile Asp Phe Ala Val Thr Lys
342             20
345 <210> SEQ ID NO: 28
346 <211> LENGTH: 13
347 <212> TYPE: PRT
348 <213> ORGANISM: Drosophila melanogaster
350 <400> SEQUENCE: 28
352 Ser Ser Asp His Ser Pro Val Leu Ile His Leu Arg Arg
353 1               5                   10
356 <210> SEQ ID NO: 29
357 <211> LENGTH: 15
358 <212> TYPE: PRT
359 <213> ORGANISM: Drosophila melanogaster
361 <400> SEQUENCE: 29
363 Met Gln Ile Ser Leu Asn Ile Val Phe Trp Asn Ala Asn Gly Leu
364 1               5                   10                  15
367 <210> SEQ ID NO: 30
368 <211> LENGTH: 10
369 <212> TYPE: PRT
370 <213> ORGANISM: Drosophila melanogaster
372 <400> SEQUENCE: 30
374 Ile Asp Ile Leu Leu Val Ser Glu Ser His
375 1               5                   10
378 <210> SEQ ID NO: 31
379 <211> LENGTH: 10
380 <212> TYPE: PRT
381 <213> ORGANISM: Drosophila melanogaster
383 <400> SEQUENCE: 31
```

VERIFICATION SUMMARY DATE: 11/21/2002
PATENT APPLICATION: **US/09/653,812B** TIME: 11:04:51

Input Set : **A:\09596-23U3 (053893-5006-03).txt**
Output Set: **N:\CRF4\11212002\I653812B.raw**